

DECLARATION UNDER 37 C.F.R. § 1.132

I, Peter Hochstein, hereby state that:

1. I am a citizen of the United States.
2. I am currently employed Relume Technologies, Oxford, MI, specializing in innovative LED light engine solutions for corporate signage, automotive, airlines, safety devices and the U.S. military.
3. I am not an inventor of the United States Patent Application No. 10/605,684 (the subject application) or United States Patent No. 6,636,237 (the parent patent).
4. I do not have any interest in the outcome of the subject application.

Background

5. I have worked in the field of video and computer related technologies for over 20 years and I am a person highly skilled in the art of such video and computer related technologies.
6. I earned a Bachelor of Science degree in Physics from Acton College in 1968.
7. In addition, I am an inventor of over 85 issued United States Patents. My inventions have included, but are not limited to, such technologies as communications between local and remote video games and display technology related to the same, optic fiber technologies, audio transducers, self-tuning antennas, rain sensors for windshield, LED devices, and many other mechanical and electrical automotive devices.
8. Further, I have been involved with encrypted, optical communication technology and solid state lighting.

The Subject Application

9. I have reviewed the subject application and the parent patent. I am familiar with the pending claims of the subject application. As I understand the subject application, claims 1, 20, 24, and 35 are in independent form.

10. The subject application, as I understand it, with reference specifically to claim 1, claims a method of retrieving information associated with an object present in a media stream by defining a user-selectable region in a layer separate from the media stream *and without accessing individual frames of the media stream*. The user-selectable region *tracks a position* of the object present in the media stream. In other words, the user-selectable region is defined without accessing individual frames and the user-selectable regions tracks a position of the object as the object moves around in the media stream. A link is defined to information associated with the object and the user-selectable region is linked in the layer to the link for the information associated with the object. Next, the user-selectable region is positioned *in the layer over the object* such that the user-selectable region *tracks the position of the object during playback of the media stream*. Again, the subject application provides that the user-selectable region tracks the position of the object which has been defined without accessing individual frames of the media stream.

11. Referring to claim 20, the subject application claims a method of providing a video signal from a provider to a user wherein a second component of a video signal is transmitted having a layer with user-selectable regions *tracking a position of objects present in the media stream* and linked to information associated with the object. The method further claims *synchronizing* the user-selectable region within the layer to a position of the object in the

media stream *without accessing individual frames of the media stream*. The user-selectable region is enabled to allow the user to select the user-selectable regions and access the information associated with the object.

12. Referring to claim 24, the subject application claims a device for storing information associated with an object present in a media stream. The device comprises a media stream with an object therein, information associated with the object, and a layer for disposition adjacent the media stream during playback. The layer has a user-selectable region *tracking a position of the object in the media stream to synchronize the user-selectable region within the layer to the position of the object in the media stream without accessing individual frames of the media stream during playback*.

13. Referring to claim 35, the subject application claims a system capable of storing and retrieving information associated with an object present in a media stream provided with a video signal from a provider. The system comprises an editor defining *a user-selectable region tracking a position of the object in the media stream without accessing individual frames of the media stream* and defining a link between the user-selectable region and information associated with the object. The system further comprises a layer disposed adjacent the media stream during playback and presenting the user-selectable region for selection by the user to access the information such that *the user-selectable region is synchronized within the layer to the position of the object in the media stream without accessing individual frames of the media stream*.

14. Generally, those of ordinary skill in the art appreciate that video is captured and played back at 30 frames per second. Thus, for a 30 minute (1800 seconds) video, there are 54,000 frames.

15. The subject application allows for quickly and seamlessly defining user-selectable regions for any object in the media stream without having to edit the individual frames of the media stream. If the user-selectable regions were defined by editing individual frames, either completely or partially, a substantial amount of resources is required. The subject application, on the other hand, developed a method and system free from accessing individual frames while still providing the user-selectable regions tracking an object in the media stream. The subject application minimizes cost associated with creating the user-selectable regions since the individual frames are not being edited which makes the technology economically feasible.

Prior Technologies

16. Prior to the subject application, as one of ordinary skill in the art, I was aware of various techniques to provide information to viewers of a media stream.

17. One method provides overlays on top of a media stream in a separate window that would provide basic information relating to the content of the media stream. Examples of this technology include interactive television guides. However, this method did not provide links to specific objects in the media stream and does not track a position of an object in the media stream.

18. Another method known to me involved editing of individual frames of the media stream and creating "hot spots" based upon the object being present in individual frames. This method requires a significant outlay of time and effort to establish the hot spots and is not practical. As mentioned above, editing frames of video for a 30 minute video would require 54,000 frames to be edited. Even if every single frame is not edited and the media stream is only

edited partially, such as one frame a second, this would still require 1,800 edits. For a 30 minute video, this is burdensome to develop.

19. Prior to the subject application, I was aware that the video industry was increasingly searching for a system and method to advertise product placement in a new medium as a result of decreases in the success of traditional advertising and increases in technologies that allowed skipping of traditional commercials. I was aware that many attempts had been made, as early as the 1980's, to develop hyperlinks in video by editing the video frame by frame and inserting the respective hyperlinks in each frame, especially in video game arts. However, prior to the subject application, the video industry has been unsuccessful to provide a system or method that would be feasible given time and budget constraints that did not rely on frame by frame editing or that merely provided a graphic overlay.

The Cited References

20. I am aware of, have read, and understand the disclosure of "Wink Communications: A Smarter Way to Watch TV", dated 08/30/2006, pages 1-13, and indicating <http://web.archive.org/web/19991012081750/http://wink.com/> (hereinafter "Wink").

21. As one of ordinary skill in the art, when considering Wink as a whole, Wink discloses a system and method for creating a form or overlay to be displayed in a viewer or screen. This overlay is very similar to traditional interactive video guides that overlay the video in response to a user pushing a button for the information. Wink provides an icon on the screen to indicate that information is available about the program. The user pushes a button and the information is retrieved.

22. Wink discloses, on page 6, that the interactive features are created by dragging objects onto a form to create the overlay. Referring to page 10, the form is designed separate from the media stream.

23. Wink discloses, on page 5, that Wink supports the use of Interactive Communicating Application Protocol (ICAP). ICAP is a compact protocol that allows for transmission in limited data bandwidth of analog broadcasts. In other words, Wink's overlay is designed to consume small amounts of bandwidth to be able to be transmitted in ICAP.

24. As one of ordinary skill in the art, Wink does not disclose, teach, or suggest defining the user-selectable region in a layer such that the user-selectable region is positioned in the layer over the object. Instead, Wink discloses that the user-selectable region is only positioned in the form and is not positioned over the object as the object moves or in the layer.

25. Thus, as one of ordinary skill in the art, without impermissibly considering the subject application, I would not have understood Wink to disclose, teach, or suggest at the time of the invention positioning the user-selectable region over the object because Wink circumvented this need by disposing the region in the form.

26. Further, Wink does not disclose, teach, or suggest to one of ordinary skill in the art that the user-selectable region is linked in the layer to a link for information about the object. Again, Wink positions the user-selectable region in the form and as such the user-selectable region is not linked in the layer to the link for the object information.

27. Thus, as one of ordinary skill in the art, without impermissibly considering the subject application, I would not have understood Wink to disclose, teach, or suggest at the time of the invention the linking of the user-selectable region in the layer because Wink circumvented this

need by the user-selectable region being present only in the form.

28. I am also aware of, have read, and understand the disclosure of “Adding Hyperlinks to Digital Television”, V. Michael Bove, Jr. et al., Proc. 140th SMPTE Technical Conference, 1998 (hereinafter “Bove”).

29. As one of ordinary skill in the art, when considering Bove as a whole, Bove discloses a system and method for creating hyperlinks in a video by accessing individual frames of the video. Bove discloses that the system will automatically create a segmentation mask for each individual frame of the video after a user identifies an object in a frame of the video. Referring to page 1, Bove states that the author scribbles on a desired object in a frame and the systems generates a segmentation mask for that frame and *following frames*. In other words, Bove creates these hyperlinks by accessing individual frames of the video.

30. With reference to page 2, Bove identifies the challenges of creating the clickable regions in every frame *manually* and the difficulty of segmenting and tracking them automatically. The solution disclosed in Bove, as set forth in the second paragraph on page 2, is generating the segmentation mask for a frame of video and continues generating the mask for following and preceding frames, i.e., frame by frame. Bove relies on the pixels in each individual frame for the segmentation mask to properly identify the object. Thus, if Bove did not edit the video frame by frame, the resultant segmentation mask would not function properly.

31. Referring to page 3, second paragraph under the heading “Segmenting Objects”, Bove states that the system classifies every pixel in every frame in the video. Further, Bove states that the author classifies pixels in a single frame and the pixels are then tracked through the remainder of the frames of the video. In the third paragraph under the heading “Segmenting

Objects”, Bove discloses that the system estimates the location of the pixels within each of the remaining frames in the video.

32. As one of ordinary skill in the art, Bove does not disclose, teach, or suggest defining user-selectable regions without accessing individual frames of the video that track a position of the object in the media stream. Instead, Bove requires that the editing be conducted frame by frame.

33. In Figure 1 on page 8, the system and method of Bove is shown whereby the user has scribbled lines in the top picture. The system creates the segmentation mask shown in the bottom picture, i.e. frame by frame. The system then generates the segmentation mask for the following and preceding frames. On page 10, Bove discloses that the segmentation mask required retraining by the user approximately every second of video. In other words, the user has to retrain the segmentation mask roughly 1800 times for a 30 minute video.

Analysis

34. As a result of my review of Wink in view of Bove, it is not be obvious to me as one of ordinary skill in the art to combine the teachings of Wink with Bove. First, the system and method disclosed in Wink merely describes an overlay that has information tied to the media stream. In other words, one skilled in the art would not be motivated to convert the overlay disclosed in Wink into user-selectable regions that track the position of the object in the media stream. Instead, Wink teaches away from developing such a system by utilizing the overlay.

35. Second, the combination of Wink with Bove has no reasonable expectation of success. As described above, Wink circumvents the issue of defining user-selectable regions that track a position of the object in the media stream by employing the generic overlay and disposing

the user-selectable regions within the overlay. Bove teaches editing the video frame by frame to locate the pixels and to create the segmentation mask. The combination of an overlay with a segmentation mask would not produce a system or method that provides user-selectable regions that tracks the position of the object without accessing individual frames of video as claimed.

36. Third, Wink discloses that ICAP is supported such that the overlays produced via Wink consume smaller amounts of bandwidth and are able to be transmitted utilizing ICAP. Bove, on the other hand, requires large amounts of processing and memory in order to handle the segmentation masks created for every second of video for many objects. Said another way, Bove would consume large amounts of bandwidth in providing the segmentation mask for individual frames for even a single object. Therefore, it would not be reasonable to expect the combination of Wink with Bove to be successful and teaches away from combining the disclosures.

37. Additionally, if even the combination of Wink and Bove were proper, the combination would not arrive at the claimed invention. As set forth above, each of the independent claims require user-selectable regions *tracking a position of objects present in the media stream* and the user-selectable regions being defined *without accessing individual frames of the media stream*. Further, the independent claims require the user-selectable region to be *synchronized* within the layer to a position of the object in the media stream again without accessing individual frames of the media stream.

38. At best, the combination of the user-selectable regions of Bove with the overlay of Wink would require the editing of the individual frames of the video to incorporate the user-selectable regions of Bove. The combination would not produce user-selectable regions that track the position of the object without accessing individual frames and that are defined without

accessing individual frames. Moreover, the combination would not produce the user-selectable region to be *synchronized* within the layer to a position of the object in the media stream again without accessing individual frames of the media stream.

39. Given that Wink teaches away from using the user-selectable regions of Bove, given that the combination has no reasonable expectation of success, and given that the combination does not arrive at the claimed invention, as one of ordinary skill in the art, there is no teaching or suggestion to combine the references. In fact, there are numerous indicia that suggest the combination is improper and each element of the claimed invention would not be shown if the combination was proper.

Conclusion

40. The subject application provides a solution that the video industry has been seeking for many years. Specifically, being able to define user-selectable regions that track a position of objects present in the media stream without accessing individual frames of the media stream. The subject application will allow a new medium of advertising to move forward. As described above, merely providing an overly has not been an adequate solution and requiring frame by frame editing has not been a solution. The subject application transcends these prior attempts and provides a solution that does not require frame by frame editing and thus provides a new solution.

41. Even in view of the cited references, as one of ordinary skill in the art, I would not have arrived at claimed system and methods of the subject application for the reasons set forth above.

42. None of the cited references disclose, teach, or suggest, alone or in combination, a

system or method of defining user-selectable regions *tracking a position of objects present in the media stream defined without accessing individual frames of the media stream.*

43. Further, none of the cited references disclose, teach, or suggest, alone or in combination, user-selectable regions *synchronized* within the layer to a position of the object in the media stream again without accessing individual frames of the media stream.

Declaration

44. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information are believed to be true, and further that these statements were made with the knowledge that willful and false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or patent issued thereon.

Respectfully submitted,

Dated Sept. 6th 2007

A handwritten signature in dark ink, appearing to read "Peter Hochstein", written over a horizontal line.

Peter Hochstein